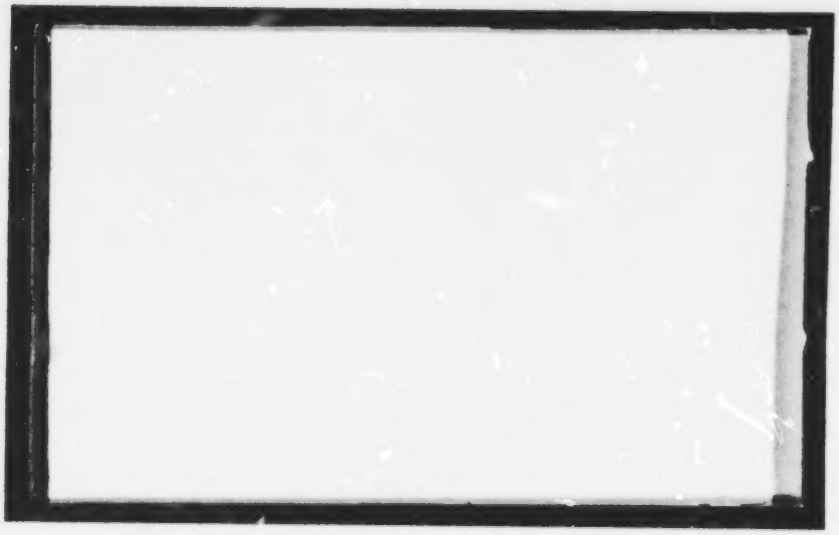
UNIVERSITÉ DE MONTRÉAL



CENTRE DE RECHERCHE SUR LES TRANSPORTS

CASE POSTALE 6128, STATION A, MONTRÉAL, QUÉBEC, H3C 3J7

A DOGIT MODEL OF TRAVEL MODE CHOICE IN MONTREAL

by

Marc J.I. Gaudry

This research was partly supported by the Research and Development Centre of Transport Canada and by the F.C.A.C. program of the *Ministère de l'Education du Québec.*

Université de Montréal

Publication # 104 - Centre de recherche sur les transports
Cahier # 7808 - Département de sciences économiques

Juillet 1978

Dépôt légal - 4e trimestre 1978

ABSTRACT

This paper compares dogit and logit specifications of a mode choice model using Montreal data on transit trips and on car vehicle trips. Box-Cox transformations of the explanatory variables are used in order not to confound the comparison with functional form specification error. The dogit specification is confidently rejected by using likelihood ratio tests. Among logit model variants, Box-Cox parameter values very different from those implied in linear and multiplicative forms of the utility functions are accepted on grounds of both statistical significance and reasonableness of estimated values of time.

CONTENTS

1. Introduction. 1

2. Model specification and interpretation 1

3. Application to Montreal 4

   A. Chosen model. 4

   B. Selected model variants 6

   C. Estimation procedure. 7

4. Results 8

   A. Tests and their significance. 8

   B. Impact of dogit parameters. 9

   C. Impact of Box-Cox transformations 12
      a. On the likelihood function 12
      b. On elasticities and the value of time. 12

   D. Other elements. 14

5. Conclusion. 14

Acknowledgements 15

References 16

Figure 1. Log likelihood surface for DOGIT U-U ($\theta,\lambda=\hat{\lambda}$) variant . . . 10

Table 1. Results for all model variants. 11

Table 2. Values of time for logit variants 14

---

1. INTRODUCTION

This paper uses urban travel mode choice data from Montreal to provide new evidence on the comparison between the logit and dogit models interpreted as market share models. The analysis uses a flexible specification of the explanatory variables in order not to confound the comparison with functional form specification error. We find, in contrast with previous results, that the logit model is preferable to the dogit model if the correct transformation of the explanatory variables is used and conclude that the sample does not support the hypothesis that consumers are captive to either mode

2. MODEL SPECIFICATION AND INTERPRETATION

The dogit discrete choice model takes the form:

$$Y_i = \frac{e^{V_i} + \theta_i \sum_j e^{V_j}}{(1 + \sum_j \theta_j) \sum_j e^{V_j}} \quad , \quad i,j = 1,\ldots,M. \qquad (1)$$

where

Y_i : probability of the i^{th} of M alternatives being chosen;

θ_i : parameter associated with the i^{th} alternative; $\theta_i \geq 0$;

V_i : function of K independent attributes X_{ik} of the i^{th} alternative.

This model, described in Gaudry and Dagenais (1977), is in general unconstrained by the independence from irrelevant alternatives (IIA) axiom because relative probabilities of choice of any two alternatives i and j in general depend on the characteristics of other alternatives. Relative probabilities of a *particular* pair of alternatives can be invariant with

respect to changes in the utility of other alternatives (if the θ_i for that pair equal zero) without imposing the invariance on *all* relative probabilities. General invariance only occurs if all θ_i equal zero: in that case the dogit collapses to the logit specification.

The θ_i can be interpreted as "income effects" which, compared to logit values, reduce the own elasticities of demand and increase the cross elasticities of demand by amounts which depend on the importance of the alternative whose characteristics are altered.

The θ_i also provide a measure of compulsive or captive behavior as can be seen by rewriting (1) as

$$Y_i = \frac{\theta_i}{1+\sum_j \theta_j} + \frac{1}{1+\sum_j \theta_j} \cdot \frac{e^{V_i}}{\sum_j e^{V_j}} \qquad (2)$$

where the first right-hand side term can be interpreted as the part of Y_i chosen to satisfy a "basic need" for commodity i and the second term is the part of Y_i determined by a logit model. If Y_i represents the share of income spent on the i^{th} commodity, the dogit is compatible with Stone's (1954) approach to consumer demand wherein part of income is used to satisfy basic needs for goods independently from their prices and the spending of discretionary income depends on prices. If Y_i represents the proportion of trips by a certain mode, the dogit reconciles in a single formula captive and free travel behavior. Ben-Akiva (1977) has recently extended this interpretation of the dogit as a model in which consumers face different subsets of the universe of alternatives and McFadden (1978) has further

shown that the dogit is a special case of mixed probability random utility models.

Model (1) can be interpreted either as a discrete choice or as a market share model. Recent tests by Gaudry and Wills (1978b) used the latter interpretation in two-alternative applications. When the number of alternatives considered equals two, the inconsistency of the dogit with IIA is not relevant but the measure of market captivity embodied in the θ_i remains of interest.

A serious problem arises when the dogit is interpreted as a market share model: the admissible region for the θ_i is constrained by the smallest observed market share of each alternative. The time-series sample used in the above-mentioned paper had a large region (due to a 50% and to a 18% minimum) but the cross-section sample had a small region (because of a 5% minimum for both alternatives): the dogit appeared to dominate the logit in the former sample but was only marginally better in the latter. The sample used in this paper has a minimum observed market share of 7% for public transportation and 10% for automobile vehicle trips, a significant increase over that of the previous cross-sectional sample.

To complete the specification of the model, it is important to determine the functional form of the V_i functions used in (1). To remove the possible bias which would arise if we compared logit and dogit models with an arbitrary choice of the functional form of the V_i functions, we shall use the following :

$$V_i = \beta_o + \sum_k \beta_k \left(\frac{x_{ik}^{\lambda_k} - 1}{\lambda_k}\right) \quad , \tag{3}$$

where the Box-Cox (1964) transformation of the non-Boolean variables includes as special cases the linear form of (3) if $\lambda = (\lambda_1, \ldots, \lambda_K) = 1$ and, given a logit format, yields the following product model if $\lambda \to 0$:

$$Y_i = \frac{\exp(\beta_o) \prod_k x_{ik}^{\beta_k}}{\sum_j \exp(\beta_o) \prod_k x_{jk}^{\beta_k}} \quad . \tag{4}$$

A constant is required in (3) to preserve the invariance of the transformation to units measurement (Schlesselman, 1971). In a 4-mode intercity logit mode choice model (Gaudry and Wills, 1978a), the use of transformation (3) led to the conclusion that form (4) gave a better explanation of observed market shares than other forms, notably the linear one. McFadden[@] has suggested that such transformations improve results because they probably represent the consumers' own perceptions of the attributes of alternatives.

3. APPLICATION TO MONTREAL

A. Chosen model

The model, obtained by substitution of (3) into (1), is

[@] McFadden made that suggestion during the NBER-NSF Conference on Decision Rules and Uncertainty at Carnegie-Mellon University in March 1978.

$$Y_i = \frac{e^{\beta_o+\sum_k \beta_k \frac{(X_{ik}^{\lambda_k}-1)}{\lambda_k}} + \theta_i \sum_j e^{\beta_o+\sum_k \beta_k \frac{(X_{jk}^{\lambda_k}-1)}{\lambda_k}}}{(1+\sum_j \theta_j) \sum_j e^{\beta_o+\sum_k \beta_k \frac{(X_{jk}^{\lambda_k}-1)}{\lambda_k}}} , \qquad (5)$$

where

Y_i = market share of mode i for origin-destination pairs for which trips are made by both modes during a typical week day morning peak;

X_{i1} = out-of-pocket fare by mode i: the transit fare is constant and the automobile fare is a constant multiple of a measure of distance for all origin-destination pairs;

X_{i2} = in-vehicle travel time by mode i: travel times are measures of actual morning peak shortest time paths in each network;

X_{i3} = wait-time by mode i: measure of actual wait-time (including transfers) by bus for each origin-destination pair; zero for automobile.

Except for the transit fare, all observations on explanatory variables are obtained from calibrated generalized time shortest-path assignment algorithms; the resulting sample values are certainly highly correlated with the true mean values which would have been obtained otherwise. The total sample consists of 113 observations pertaining to 25 origin and destination zones in Montreal in November 1970.

B. Selected model variants

In the absence of restrictions, there are 9 distinct coefficients in (5):

$$\theta = (\theta_1 , \theta_2)$$

$$\lambda = (\lambda_1 , \lambda_2 , \lambda_3)$$

$$\beta = (\beta_0 , \beta_1 , \beta_2 , \beta_3)$$

We shall select for estimation the following set of dogit and logit variants defined by varying the restrictions on θ and λ:

a) DOGIT U-U: dogit unconstrained-unconstrained

The first and second letters of the code name for this model respectively mean that θ and λ are unconstrained.

b) DOGIT CE-U: dogit constrained equal-unconstrained

This model is defined by imposing the equality constraint $\theta = (\theta_1 = \theta_2)$ and leaving $\lambda = (\lambda_1 , \lambda_2 , \lambda_3)$ unrestricted.

c) DOGIT U-CE: dogit unconstrained-constrained equal

The reverse of (b): θ is unrestricted but the elements of λ are constrained equal.

d) DOGIT U-CM: dogit unconstrained-constrained multiplicative

As in (c), but constraining $\lambda \to 0$, which defines the dogit version of multiplicative model (4).

e) DOGIT U-CL: dogit unconstrained-constrained linear

As in (c), but constraining $\lambda = 1$, which corresponds to a linear exponent dogit specification.

b') LOGIT-U: logit unconstrained

As in (b), but requiring $\theta = 0$.

c') LOGIT-CE: logit constrained equal

As in (c), but requiring $\theta = 0$.

d') LOGIT-CM: logit constrained multiplicative

As in (d), but requiring $\theta = 0$, which defines the product share model (4).

e') LOGIT-CL: logit constrained linear

As in (c), but requiring $\theta = 0$, which corresponds to the standard linear exponent logit specification.

C. Estimation procedure

Taking logarithms of the ratios of market shares Y_1 and $Y_2 = 1 - Y_1$ yields the following form of (5):

$$\ell n \left[\frac{Y_1 - (\theta_1 / 1+\theta_1+\theta_2)}{Y_2 - (\theta_2 / 1+\theta_1+\theta_2)} \right] = \beta_0 + \sum_k \beta_k \left(\frac{(X_{1k}^{\lambda_k} - 1)}{\lambda_k} - \frac{(X_{2k}^{\lambda_k} - 1)}{\lambda_k} \right) + u \quad , \tag{7}$$

where the error term is assumed to be normally and identically distributed with a mean equal to zero and a constant variance σ^2 and where the admissible region for θ_1 and θ_2 is

$$[\theta_1 / 1+\theta_1+\theta_2 < \min.(Y_1)] \text{ and } [\theta_2 / 1+\theta_1+\theta_2 < \min.(1-Y_1)] \quad . \tag{8}$$

Given these assumptions, the likelihood of the original observations on market share Y_1 is

$$L(Y_1;\beta,\sigma^2,\lambda,\theta) = (2\pi\sigma^2)^{\frac{-N}{2}} \cdot \exp\left(\frac{-u'u}{2\sigma^2}\right) \cdot |J(\theta;Y_1)|, \tag{9}$$

where $|J(\theta;Y_1)|$ is the absolute value of the Jacobian of the transformation from the dependent variable in (7) to the actually observed Y_1, or

$$|J(\theta;Y_1)| = \prod_{t=1}^{N} \{(1+\theta_1+\theta_2)\left[Y_{1_t} - \frac{\theta_1}{1+\theta_1+\theta_2}\right]\left[1-Y_{1_t} - \frac{\theta_2}{1+\theta_1+\theta_2}\right]\}^{-1},$$

$$= \prod_{t=1}^{N} J_t, \tag{10}$$

where the subscript t denotes observations $1,\ldots,N$. It is convenient to write the log concentrated likelihood function

$$L(Y_1;\lambda,\theta) = -\frac{N}{2}[\ell n(2\pi)+1] - \frac{N}{2}\ell n\left[\frac{u'u}{N}\right] + \sum_{t=1}^{N} \ell n\, J_t \tag{11}$$

which can be maximized by nonlinear methods or more simply by scanning. We adopted the latter procedure and performed a grid search over values of θ and λ.

4. RESULTS

A. Tests and their significance

A natural test of the results as their pertain to the parameters θ and λ is to compare the maximized likelihood by means of the likelihood ratio test. The test statistic for comparisons of the log likelihood values is:

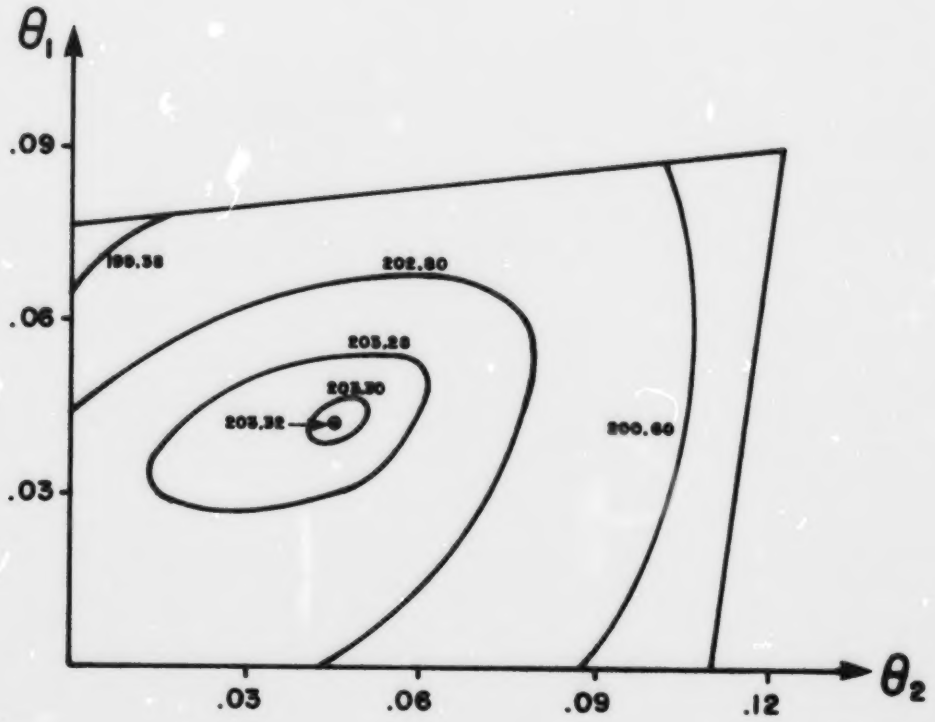
$$[L(\Omega) - L(\omega)] > \frac{1}{2}\chi^2_{\alpha}\,(\text{D.F.}), \tag{12}$$

where Ω is the parameter space under a maintained hypothesis, ω is the

restricted space generated by a null hypothesis and the right-hand side expression denotes the chi-square distribution with given degrees of freedom and significance level α. For given values of θ and λ, Student's t distribution can be used to test for the statistical significance of the β_k parameters. Performing such conditional tests on the β_k and their standard errors is a weaker procedure than performing joint tests on these parameters and on θ and λ simultaneously: the resulting significance levels are probably overestimated.

B. Impact of dogit parameters

The general impact of dogit parameters θ_1 and θ_2 is clearly shown in Figure 1 which exhibits the shape of the log likelihood function for various values of these parameters, given the optimal values $\hat{\lambda}_1$, $\hat{\lambda}_2$, and $\hat{\lambda}_3$ estimated in model variant DOGIT U-U. The unique maximum occurs approximately at the center of the admissible region. The surface is asymmetric and it is immediately apparent that dogit parameters have a weak impact on the likelihood of the observations and do not make the logit solution (at the origin) significantly less desirable. This impression in confirmed by an analysis of the results shown in Table 1. A pairwise comparison of dogit and logit variants for identically constrained λ vectors (U-U vs -U, U-CE vs -CE, U-CM vs -CM, U-CL vs -CL) shows, given the 2 degrees of freedom required, that the logit model can never be reasonably rejected because the value of the test statistic is in all cases less than 1. This is confirmed in the R-squared values, computed at the mean of variables transformed as in (7), which are on the whole barely affected by the addition of θ parameters. It is not surprising to find also that the implied values of



LOG LIKELIHOOD SURFACE FOR DOGIT U-U (θ, $\lambda = \hat{\lambda}$) VARIANT

FIGURE 1

the captive market shares $(\theta_i/(1+\sum_j \theta_j)$ are either 0% or approximately 4%. We can safely reject the hypothesis that consumers are captive to either of the modes and, using Ockham's razor, accept the logit model for this data set.

TABLE 1. RESULTS FOR ALL MODEL VARIANTS *

	DOGIT					LOGIT			
	U-U	CE-U	U-CE	U-CM	U-CL	-U	-CE	-CM	-CL
θ_1 transit	.0422	.0431	.0414	.000	.000	0.0	0.0	0.0	0.0
θ_2 auto	.0449	.0431	.0468	.042	.040	0.0	0.0	0.0	0.0
λ_k fare	4.496	4.587	3.855	0.0	1.0	4.212	3.570	0.0	1.0
λ_k travel-time	0.557	0.542	3.855	0.0	1.0	0.655	3.570	0.0	1.0
λ_k wait-time	4.619	4.672	3.855	0.0	1.0	4.215	3.570	0.0	1.0
fare coeff.	-9.102	-9.298	-11.489	-0.11	-1.32	-7.37	-9.43	-0.10	-1.23
fare t-stat.	(-2.17)	(-2.17)	(-1.67)	(-0.71)	(-2.04)	(-2.13)	(-1.78)	(-0.70)	(-2.04)
fare elast.**	-.02/.02	-.02/.02	-.04/.04	-.05/.05	-.17/.17	-.02/.02	-.05/.05	-.05/.05	-.17/.17
tr.-t. coeff.	-0.81	-0.085	-0.4×10^{-6}	-0.48	-0.024	-0.056	-0.1×10^{-5}	-0.45	-0.02
tr.-t. t-stat.	(-1.32)	(-1.32)	(-0.81)	(-1.43)	(-1.59)	(-1.45)	(-1.00)	(-1.44)	-1.57
tr.-t. elast.**	-.20/.15	-.20/.15	-.04/.01	-.21/.21	-.27/.15	-.21/.15	-.06/.01	-.21/.21	-.26/.15
wait-t. coeff.	-0.3×10^{-4}	-0.3×10^{-4}	-0.2×10^{-3}	-0.56	-0.11	-0.5×10^{-4}	-0.4×10^{-3}	0.54	-0.11
wait-t. t-stat.	(-5.17)	(-5.13)	(-4.93)	(-2.97)	(-3.83)	(-4.90)	(-4.69)	(-3.04)	(-3.93)
wait-t. elast.**	-.09/n.a.	-.09/n.a.	-.14/n.a.	-.25/n.a.	-.33/n.a.	-.09/n.a.	-.14/n.a.	-.26/n.a.	-.34/n.a.
constant coeff.	0.540	0.532	0.38	1.53	1.14	0.48	0.35	1.35	0.99
constant t-stat.	(2.84)	(2.78)	(3.86)	(4.44)	(5.45)	(2.99)	(3.90)	(4.22)	(5.02)
$L(\Omega)$	203.320	203.319	202.810	196.349	200.366	202.973	202.460	195.968	200.034
$L(\Omega) - L(\omega)$	0.0	0.001	0.510	6.971	2.954	0.397	0.86	7.352	3.286
R^2	.24	.24	.23	.12	.18	.23	.22	.12	.18
D.F.	0	1	2	3	3	2	4	5	5

* U = unconstrained, C = constrained, E = equal, M = multiplicative, L = linear.

** The own and cross elasticities for transit (computed at the mean).

C. Impact of Box-Cox transformations

a. On the likelihood function

Not any logit model should however be accepted. There are strong grounds for choosing among logit variants (and, had the dogit model been the only estimated model, among dogit variants). The multiplicative model is absolutely dominated by the linear model; the latter is rejected in favor of the constrained equal model at the 3% level (the value of the test statistic is 2.426 and there is one degree of freedom). Sacrificing 2 degrees of freedom to permit, in the unconstrained model, unequal elements of λ does not increase the value of the likelihood sufficiently to reject this LOGIT-CE model with a high degree of confidence.

b. On elasticities and the value of time

There may be other than statistical grounds for choosing between LOGIT-CE and LOGIT-CL variants. The reasonableness of demand elasticities is often used as a practical measure because models with different demand elasticities have different implications for policy decisions based on the responses which elasticities imply. The formulas for own and cross elasticities in the dogit model are respectively:

$$\eta_{ii} = \beta_k X_{ik}^{\lambda_k} \left[\frac{(1+\theta_i)e^{V_i}}{e^{V_i}+\theta_i \sum_j e^{V_j}} - \frac{e^{V_i}}{\sum_j e^{V_j}} \right]$$

and

$$\eta_{ij} = \beta_k X_{jk}^{\lambda_k} \left[\frac{\theta_i e^{V_j}}{e^{V_i}+\theta_i \sum_j e^{V_j}} - \frac{e^{V_j}}{\sum_j e^{V_j}} \right] . \qquad (13)$$

On the basis of elasticity values (obtained in the logit by setting the θ_i of (13) equal to zero), there appears to be little to choose between CE and CL logit variants: the transit demand own elasticities shown in Table 1 are quite reasonable in that own price elasticities are lower than own travel-time elasticities and the latter generally lower than own wait-time elasticities; it is also reasonable that transit cross elasticities be equal to or lower than own elasticities in both variants. But this evidence is not conclusive and other evidence should be considered.

It is useful to consider time prices derived from the parameter estimates in standard fashion by considering the ratios of derivatives of the demand function with respect to time and price variables, a measure which is well known (DeSerpa, 1971) to assume that the rates of change of the price of time with respect to income and the money cost of the trip are either zero or equal and of opposite signs. The Box-Cox transformation naturally affects the β coefficients and a "better" model should yield more reasonable time values. Time prices for all logit variants are shown in Table 2. The average weekly hourly wage rate in manufacturing in 1970 was approximately \$3.50 per hour: the results for the all variants imply that consumers dislike waiting for the bus more than working. The relative valuations of travel time by transit and car users support the likelihood ratio rest results and provide further evidence that the CE model with $\hat{\lambda} = 3.57$ should be preferred to the "second best" CL model because one expects the price of travel time by car to be much higher than the price of travel time by transit.

TABLE 2

VALUES OF TIME FOR LOGIT VARIANTS

			U	CE	CM	CL
Transit users	Price of travel-time	$/hr	$7.43	$0.85	$2.98	$1.08
	Price of wait-time	$/hr	$12.40	$7.71	$14.33	$5.50
Car users	Price of travel-time	$/hr	$9.24	$2.40	$5.73	$1.11
	Price of wait-time	$/hr	n.a.	n.a.	n.a.	n.a.

D. Other elements

Socio-economic factors were not included among the explanatory variables because it was felt that the introduction of alternative-specific λ_{ik} parameters suggested elsewhere (Gaudry, 1978) and required in (3) for the socioeconomic variables not to disappear in (7) would have considerably increased the computational burden or led to arbitrary simple specifications (such as associating income with one mode and setting its value to zero for the other). It was also felt that the results with basic variables which vary across alternatives, such as price and time, should be reasonably robust. If one believes that the excluded variables are very important, the results obtained are clearly biased and do not strongly establish that, for this data set, the LOGIT-CE model is the most acceptable model.

5. CONCLUSION

We have shown, in a model where the choice of travel mode in Montreal is assumed to depend solely on the price and service characteristics of the competing alternatives, that there existed no statistically significant

consumer captivity to either transit or the car (drivers' trips). It may be that an 8% or 10% lower bound on observed market shares of alternatives is not sufficient to insure that the dogit model is even marginally "better" than the logit model: an analysis of the complex trade-off between lowest sample shares and the size distribution of shares would be required to comment precisely on this problem of the dogit market share model. We have also found that the functional form parameters of the utility functions associated with both dogit and logit models had more importance than the parameters which make us choose between these two model types. Both statistical reasons and the values of time derived from various logit variants supported Box-Cox parameter values very different from those implied by linear or multiplicative forms of the utility functions.

ACKNOWLEDGEMENTS

The research on which this article is based was supported partly by the Research and Development Centre of Transport Canada and by the F.C.A.C. program of the *Ministère de l'Éducation du Québec*. The Montreal Urban Community Transit Commission supplied unpublished data. Danielle Bilodeau wrote the Fortran program and Sylvie Préfontaine performed computations. Any misuse of these resources is solely the author's responsibility.

REFERENCES

Ben-Akiva, M. (1977) Choice models with simple choice set generating processes. Working paper. Department of civil engineering, M.I.T.

Box, G.E.P. and Cox, D.R. (1964) An analysis of transformations. Journal of the Royal Statistical Society, Series B, 211-243.

DeSerpa, A.C. (1971) A theory of the economics of time. The Economic Journal 81, 828-846.

Gaudry, M.I.J. (1978) A generalized logit mode choice model. Publication # 98, Centre de recherche sur les transports, Université de Montréal.

Gaudry, M.I.J. et Dagenais, M.G. (1977) The dogit model. Publication # 82, Centre de recherche sur les transports, Université de Montréal. Forthcoming in Transportation Research.

Gaudry, M.I.J. and Wills, M.J. (1978a) Estimating the functional form of travel demand models. Transportation Research 12, 257-289.

Gaudry, M.I.J. and Wills, M.J. (1978b) Testing the dogit model with aggregate time-series and cross-sectional travel data. Publication # 94, Centre de recherche sur les transports, Université de Montréal. Forthcoming in Transportation Research.

McFadden, D. (1978) Econometric models for probabilistic choice. Paper presented at the NBER-NSF Conference on Decision Rules and Uncertainty, Carnegie-Mellon University. Mimeographed. p. 19.

Schlesselman, J. (1971) Power families: a note on the Box and Cox transformation. Journal of the Royal Statistical Society, Series B. Vol. 33.

Stone, R. (1954) Linear expenditure systems and demand analysis: an application to the pattern of British demand. The Economic Journal 64, 511-527.